Update on response fit

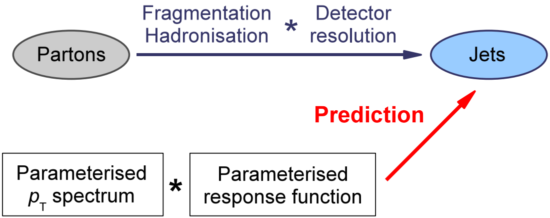
UHH CMS SUSY Meeting

Matthias Schröder

February 11th, 2010

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Concept of the response fit method



- In each event i, probability density of the dijet configuration p_T^1, p_T^2 is

$$\mathcal{P}_i^{1,2} \propto \int_0^\infty \mathrm{d}p_T^{\mathrm{true}}\, f_{\mathbf{b}}(p_T^{\mathrm{true}}) \cdot r_{\mathbf{b}}(p_T^1/p_T^{\mathrm{true}}) \cdot r_{\mathbf{b}}(p_T^2/p_T^{\mathrm{true}})$$

- $f_{\mathbf{b}}$ is the probability density function (pdf) of p_T^{true}
- $r_{\mathbf{b}}$ is the response pdf

- Likelihood $\tilde{\mathcal{L}}(\mathbf{b}) = \prod_{i=0}^{N_{\mathrm{evt}}} \mathcal{P}_i^{1,2}$ maximal for correct parameter values $\mathbf{b}$

Current studies

- New response parameterisation with Crystal Ball Function

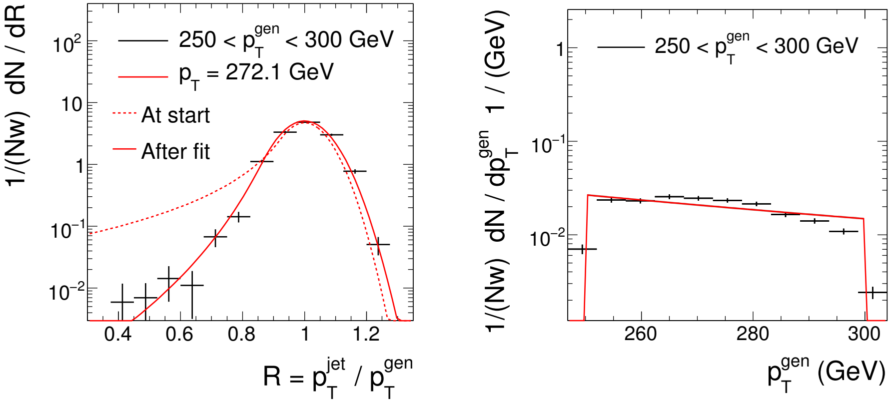
$$f(x;\bar{x},\sigma,\alpha,n) = N \begin{cases} e^{\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2} & \text{if } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A(\alpha,n)\cdot\left(B(\alpha,n) - \frac{x-\bar{x}}{\sigma}\right)^{-n} & \text{if } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases}$$

1. Fits in different $p_{\mathrm{T}}^{\mathrm{gen}}$ bins
   - Good performance, consistent HT and MHT predictions
   - Small statistical errors, but how meaningful ($\chi^2 \rightarrow \mathcal{L}$)?
   - Large anti-correlation between α and n
2. Interpolation of fitted parameters over larger p_{T} range
   - Technically challenging due to complex integrals
   - Under investigation
3. Fits in different $p_{\mathrm{T}}^{\mathrm{dijet}}$ bins
   - Large anti-correlation between α and n
   - Run-away parameter n

Event selection

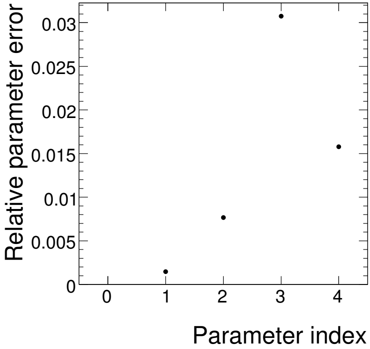
- A few results shown in the following
- `/QCDFlat_Pt15to3000/Summer09-MC_31X_V9_7TeV-v1/GEN-SIM-RECO`
- Weighting $\propto \hat{p}_{\mathrm{T}}^{-4.5}$ for real QCD spectrum
- L2L3 corrected AK5 `calojets`
- The usual dijet event selection
  - 2 jets leading in (corrected calo) p_{T} with $|\eta| < 0.8$
  - $p_{\mathrm{T}}^{3}/p_{\mathrm{T}}^{\mathrm{dijet}} < 0.1$ or $p_{\mathrm{T}}^{3} < 2$ GeV, where $p_{\mathrm{T}}^{\mathrm{dijet}} = \frac{p_{\mathrm{T}}^{1}+p_{\mathrm{T}}^{2}}{2}$
  - $|\Delta\phi^{1,2}| > 2.7$
  - $0.07 < p_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{1,2} < 0.95$
  - $\Delta R(\texttt{jet}, \texttt{genjet}) < 0.1$ (important for validation)
- Both jets in central region $|\eta| < 0.8$
- Eventually different η regions "and cross-terms"

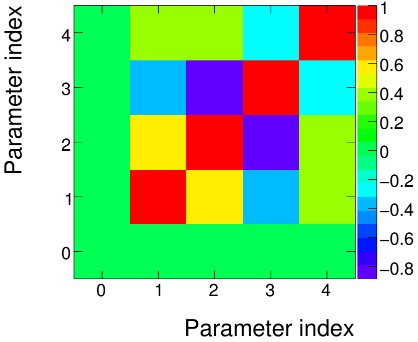
Fitted response and spectrum for $250 < p_T^{gen} < 300$ GeV



- Spectrum parameterised $\propto (p_T^{gen})^{-m}$
- Propagated errors not visible

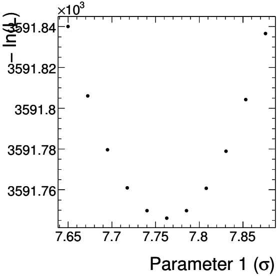
Errors and correlations for $250 < p_T^{gen} < 300$ GeV

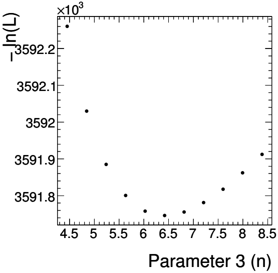


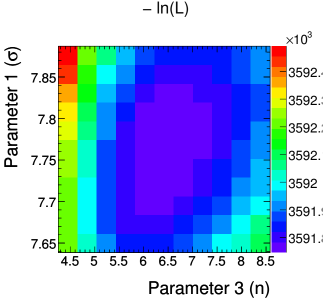


- Parameter 0 (mean of Gaussian i.e. JES) is fixed during fit
- Relative errors on fitted parameters $\leq 3\%$
- Correlation between
  - Exponent of tail n (3) and start of tail α (2)
  - Gaussian width σ (1) and start of tail α (2)
  - Exponent m of spectrum (4) with σ (1) and α (2)

Stability of minima for $250 < p_T^{gen} < 300$ GeV

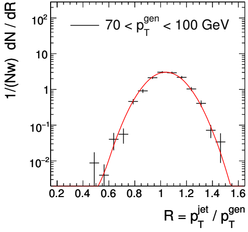


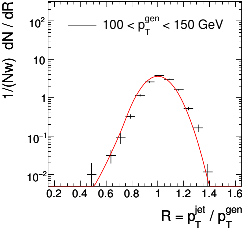


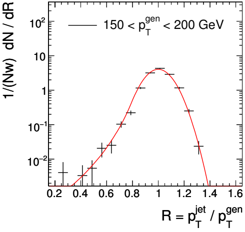


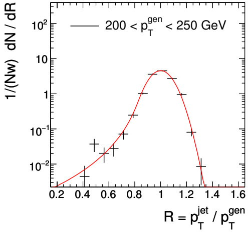
- Parameter variation 5σ around fitted value
- Not too far from parabola
- $\Delta \ln(P) < 1\%$ (large Gaussian contribution in r, good start values for σ)
- Statistical errors $\approx 68\%$ CL

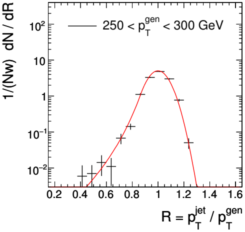
Fitted response in different p_T^{gen} bins (1)

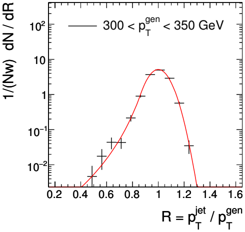




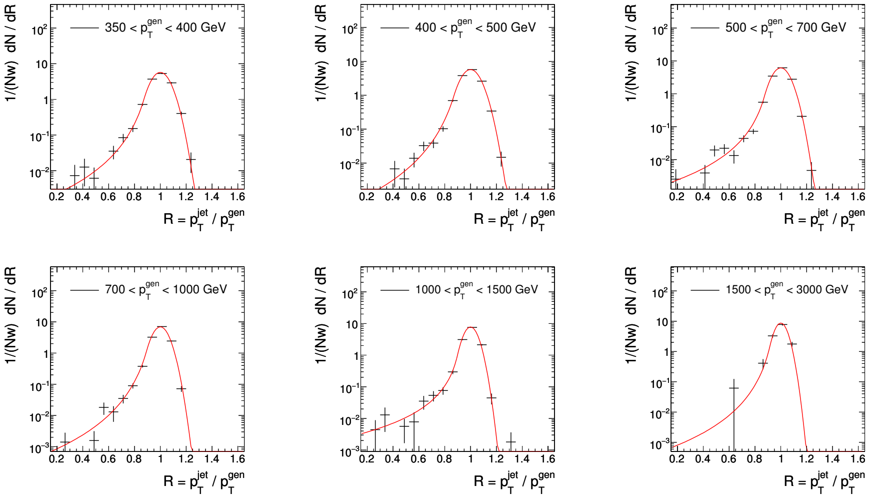




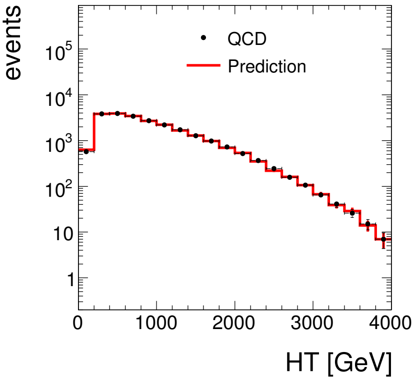


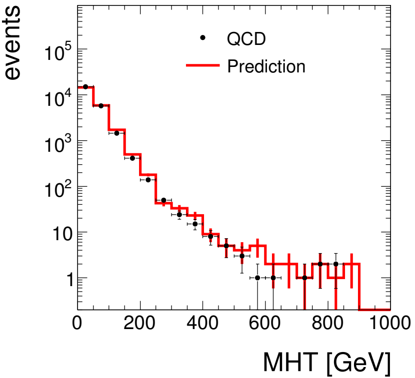


Fitted response in different p_T^{gen} bins (2)



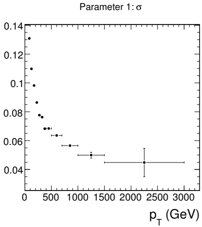
Closure: MHT prediction

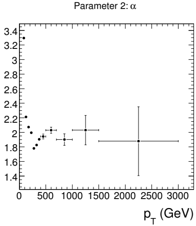


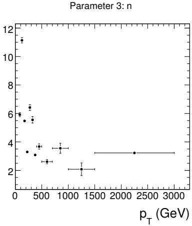


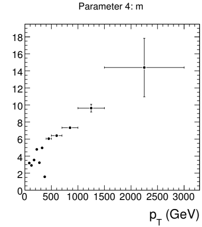
- Same dijet event selection as for the fit ($70 < p_T^{gen} < 3000$ GeV)
- Two leading `genjets` were smeared
- HT and MHT from two leading corrected `calojets` and smeared jets

Interpolation of fitted parameters





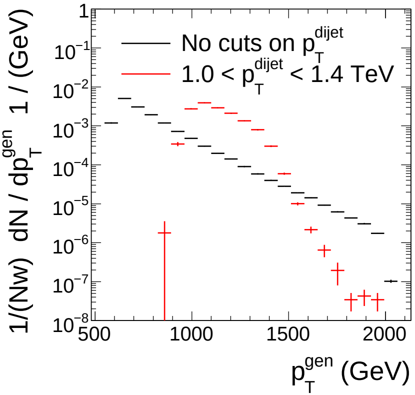




- Visible trends but significant fluct.

- Fits with some interpolation functions instable
  - Major contribution from low p_T events
  - Pure technical problem (bugs)?
  - Choice of interpolation function?
  - Fundamental problem of method (correlation between σ and spectrum)?
- Closer investigation using ToyMC

Different approach: fits in p_T^{dijet} bins

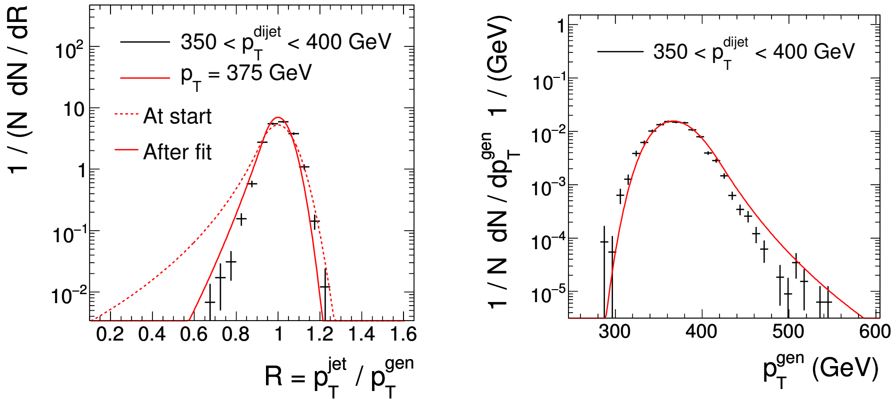


- Bin: $p_{T,min}^{dijet} < p_T^{dijet} < p_{T,max}^{dijet}$
- Response pdf r
- Pure spectrum $f_0 \propto 1/(p_T^{true})^m$

Inclusion of p_T^{dijet} cuts into pdf of p_T^{true}

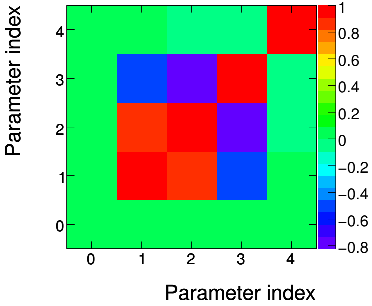
$$f(p_T^{true}) \propto f_0(p_T^{true}) \int_{p_{T,min}^{dijet}}^{p_{T,max}^{dijet}} dx \, \frac{r(x/p_T^{true})p_T^{true}}{\sqrt{2}}$$

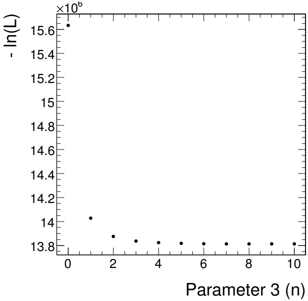
Fitted response and spectrum for $350 < p_T^{\text{dijet}} < 400$ GeV



- Parameter erros $\leq 8\%$, propagated errors not visible
- Contribution of fitted tail too large

Instabilities





- Strong correlation between
  - Exponent of tail n (3) and start of tail α (2)
  - Gaussian width σ (1) and start of tail α (2)
- In some p_T^{dijet} bins no minimum for reasonable values of n
  - Run-away $n \to \infty$, compensated by smaller α
- Problem with precision of numerical integration of $f(p_T^{true})$?
- Implementation of Crystal Ball function suboptimal at some points